

QGIS Application - Bug report #20025

GDAL/OGR Dissolve algorithm not properly working with point/multipoint layers

2018-10-05 02:02 PM - Andrea Giudiceandrea

Status: Closed	
Priority: Normal	
Assignee: Alexander Bruy	
Category: Processing/OGR	
Affected QGIS version: 3.4.4	Regression?: No
Operating System:	Easy fix?: No
Pull Request or Patch applied: Yes	Resolution: fixed/implemented
Crashes QGIS or corrupts data: No	Copied to github as #: 27847

Description

If the input layer is a point/multipoint layer and the output layer is set to "Save to temporary file" or is a shapefile, then the GDAL/OGR Dissolve algorithm fails (in most of the cases) to properly generate the correct dissolved layer and the log shows the following error

GDAL command output:

ERROR 1: Attempt to write non-multipoint (POINT) geometry to multipoint shapefile.

ERROR 1: Unable to write feature 1 from layer SELECT.

ERROR 1: Terminating translation prematurely after failed

translation from sql statement.

This is due to the fact that "OGR shape driver doesn't handle writing points to a multipoint shp" (see <https://trac.osgeo.org/gdal/ticket/3677>).

A simple workaround is to use

```
SELECT ST_Multi(ST_Union({}))
```

instead of

```
SELECT ST_Union({})
```

in <https://github.com/ggis/QGIS/blob/master/python/plugins/processing/algs/gdal/Dissolve.py#L168-L172>

```
if self.parameterAsBool(parameters, self.KEEP_ATTRIBUTES, context):
```

```
    sql = "SELECT ST_Union({}) AS {{{}} FROM '{}{}'.format(geometry, geometry, other_fields, params, layerName, group_by)
```

```
    else:
```

```
        sql = "SELECT ST_Union({}) AS {{{}} FROM '{}{}'.format(geometry, geometry, ' ' + fieldName if fieldName else ' ', params, layerName, group_by)
```

if the input layer is of point or multipoint geometry type.

Associated revisions

Revision 32f6034b - 2019-01-24 11:57 PM - Alexander Bruy

[processing][needs-docs] force multipart output from GDAL-based dissolve algorithm (fix #20025)

Revision f8893d76 - 2019-01-25 01:34 PM - Alexander Bruy

[processing][needs-docs] force multipart output from GDAL-based dissolve algorithm (fix #20025)

(cherry picked from commit 32f6034be708b305ed4e19b4f6ade1a8b409993b)

History

#1 - 2018-10-05 02:52 PM - Giovanni Manghi

- Assignee set to Giovanni Manghi

Assigning to myself unless you have a patch in the works(?). This probably affects also the buffer tool (that has a dissolve option).

#2 - 2018-10-05 04:15 PM - Andrea Giudiceandrea

Giovanni Manghi wrote:

| *Assigning to myself unless you have a patch in the works(?)*.

I'm not working on a patch. Thanks for your commitment.

| *This probably affects also the buffer tool (that has a dissolve option)*.

ST_Union() is used in Buffer.py and in OneSideBuffer.py but they are not affected because the output layer for these algs is of polygon geometry type and the issue doesn't occur for line/multiline or polygon/multipolygon shapefile output layers.

#3 - 2018-10-28 01:56 PM - Andrea Giudiceandrea

- Affected QGIS version changed from 3.3(master) to 3.4.0

Confirmed on 3.4.0

#4 - 2019-01-22 10:58 AM - Alexander Bruy

Can you share small dataset allowing to reproduce issue?

#5 - 2019-01-22 10:18 PM - Giovanni Manghi

- Status changed from Open to Feedback

#6 - 2019-01-22 11:03 PM - Andrea Giudiceandrea

- File testpoints.zip added

Alexander Bruy wrote:

| Can you share small dataset allowing to reproduce issue?

The GDAL/OGR Dissolve algorithm (Dissolve field set to TestField - output set to "Save to temporary file" or to a shapefile) fails with the attached testpoints.zip zipped Shapefile that contains the following 6 points:

	id	TestField
0	1	AAA
1	2	BBB
2	3	AAA
3	4	AAA
4	5	CCC
5	6	BBB

The algorithm log window shows:

Processing algorithm...

Algorithm 'Dissolve' starting...

Input parameters:

```
{ 'COMPUTE_AREA' : False, 'COMPUTE_STATISTICS' : False, 'COUNT_FEATURES' : False, 'EXPLODE_COLLECTIONS' : False, 'FIELD' : 'TestField', 'GEOMETRY' : 'geometry', 'INPUT' : 'C:\\testdissolvepoint\\testpoints.shp', 'KEEP_ATTRIBUTES' : False, 'OPTIONS' : '', 'OUTPUT' : 'C:/testdissolvepoint/testdissolvedpoints.shp', 'STATISTICS_ATTRIBUTE' : None }
```

GDAL command:

```
ogr2ogr C:/testdissolvepoint/testdissolvedpoints.shp C:\\testdissolvepoint\\testpoints.shp -dialect sqlite -sql "SELECT ST_Union(geometry) AS geometry, TestField FROM 'testpoints' GROUP BY TestField" -f "ESRI Shapefile"
```

GDAL command output:

ERROR 1: Attempt to write non-multipoint (POINT) geometry to multipoint shapefile.

ERROR 1: Unable to write feature 2 from layer SELECT.

ERROR 1: Terminating translation prematurely after failed

translation from sql statement.

Execution completed in 1.96 seconds

Results:

```
{'OUTPUT': <QgsProcessingOutputLayerDefinition {'sink':C:/testdissolvepoint/testdissolvedpoints.shp, 'createOptions': {'fileEncoding': 'System'}}>}
```

Loading resulting layers

Algorithm 'Dissolve' finished

The resulting output Shapefile contains only the following 2 multipoint features

	TestField
0	AAA
1	BBB

instead of 3 multipoint features:

```
TestField
0 AAA
1 BBB
2 CCC
```

#7 - 2019-01-22 11:16 PM - Giovanni Manghi

- Status changed from Feedback to Open
- Affected QGIS version changed from 3.4.0 to 3.4.4

Andrea Giudiceandrea wrote:

Alexander Bruy wrote:

Can you share small dataset allowing to reproduce issue?

The GDAL/OGR Dissolve algorithm (Dissolve field set to TestField - output set to "Save to temporary file" or to a shapefile) fails with the attached testpoints.zip zipped Shapefile that contains the following 6 points:

this tool is ogr2ogr based, specifically ogr2ogr is used to launch a spatial query, in this case:

```
SELECT ST_Union(geometry) AS geometry, TestField FROM 'testpoints' GROUP BY TestField
```

we must modify it and add a "-nlt MULTI*" parameter.

#8 - 2019-01-23 06:35 PM - Alexander Bruy

Giovanni Manghi wrote:

this tool is ogr2ogr based, specifically ogr2ogr is used to launch a spatial query, in this case:

```
SELECT ST_Union(geometry) AS geometry, TestField FROM 'testpoints' GROUP BY TestField
```

we must modify it and add a "-nlt MULTI" parameter.*

I see, but works fine even without it with my test data and QGIS test data. That's why I asked for some dataset.

Also I'm not sure if always casting output to multigeometry is correct. As I understand, dissolve also can produce singleparts. Maybe better to implement some logic and produce multiparts only if input layer is multipart?

#9 - 2019-01-23 11:42 PM - Andrea Giudiceandrea

Alexander Bruy wrote:

Also I'm not sure if always casting output to multigeometry is correct. As I understand, dissolve also can produce singleparts. Maybe better to

| *implement some logic and produce multipart only if input layer is multipart?*

Note that the provided "testpoints" dataset is a simple point shapefile, not a multipoint.

If trivially dissolved by the "id" field, the output shapefile could have a simple point geometry type.

But if dissolved by the "TestField" field, the output will contain 2 multipart point features and 1 singlepart point feature: in this case, the only way to store the output in a shapefile, is to force ogr2ogr to always create multipart point features and use a multipoint shapefile (because point and multipoint features cannot be stored together in a point shapefile nor in a multipoint shapefile).

As previously said, the issue doesn't occur for mixed singlepart/multipart features in a PolyLine or in a Polygon shapefile.

So I think that, only for point/multipoint input layer, the GDAL/OGR Dissolve algorithm should be fixed in order to always generate multipoint features as output, at least when the output is set to "Save to temporary file" (the default option) or is a shapefile.

We can use

```
SELECT ST_Multi(ST_Union({})) instead of SELECT ST_Union({}) in the sql string passed to ogr2ogr
```

as proposed by me,

or we can add "-nlt MULTIPPOINT" to the ogr2ogr command

as proposed by Giovanni.

EDIT: also "-nlt PROMOTE_TO_MULTI"

#10 - 2019-01-24 02:26 AM - Niyall Dawson

I think always upgrading to multipart is the correct choice here -- that's what the native QGIS dissolve algorithm does too.

#11 - 2019-01-24 06:35 PM - Alexander Bruy

- *Pull Request or Patch supplied changed from No to Yes*
- *Status changed from Open to In Progress*
- *Assignee changed from Giovanni Manghi to Alexander Bruy*

PR <https://github.com/qgis/QGIS/pull/8972>

#12 - 2019-01-24 11:56 PM - Alexander Bruy

- *% Done changed from 0 to 100*
- *Status changed from In Progress to Closed*

Applied in changeset commit:qgis|32f6034be708b305ed4e19b4f6ade1a8b409993b.

#13 - 2019-01-25 07:08 AM - Alexander Bruy

- *Resolution set to fixed/implemented*

Files

testpoints.zip

1.39 KB

2019-01-22

Andrea Giudiceandrea